Statistics 2 - PMCC - Solutions

Section 1

1. X = no. days with "light" windspeed

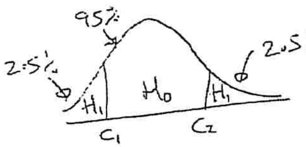
a) $X \sim Bin(20, 0.9)$

i) $P(X < 16) = P(X \le 15) = \underline{0.0432}$ (4dp)

ii) $P(X > 15) = 1 - P(X \le 15) = 1 - 0.0432 = \underline{0.9568}$ (4dp)

b) i) $H_0: p = 0.9$

$H_1: p \neq 0.9$



$P(X \le c_1) = 0.025$

$P(X \le 14) = 0.0112$

$P(X \le 15) = 0.0431$ } c_1 between 14 and 15

$P(X \geq c_2) = 0.025$

$P(X \le c_2 - 1) = 0.975$

$P(X \le 19) = 0.8784$ } $c_2 - 1$ between 19 and 20

$P(X \le 20) = 1$ } c_2 between 20 and 21

But $n = 20 \Rightarrow$ no H_1 region on the right

i.e. $\underline{\text{critical region is } x \le 14}$

ii) $x = 16 \Rightarrow$ $\underline{\text{Accept } H_0, \text{ Reject } H_1}$ i.e. insufficient evidence that Michelle's calculations are wrong at 5% level.

2. a) $P(S_1 \cap S_2 \cap S_3') = 0.2 \times 0.3 \times 0.9 = \underline{0.054}$

b) $P(S_1 \cap S_2 \cap S_3') + P(S_1 \cap S_2' \cap S_3) + P(S_1' \cap S_2 \cap S_3)$

$= 0.2 \times 0.3 \times 0.9 + 0.2 \times 0.7 \times 0.1 + 0.8 \times 0.3 \times 0.1$

$= 0.054 + 0.014 + 0.024 = \underline{0.092}$

c) $P(S_1 / \text{Stop at } 2) = \dfrac{P(S_1 \cap S_2 \cap S_3') + P(S_1 \cap S_2' \cap S_3)}{P(\text{Stop at } 2)}$

$= \dfrac{0.054 + 0.014}{0.092} = \underline{0.7391}$ (4dp)

3.

	A	A'	
B	0.1	0.4	0.5
B'	0.2	0.3	$x = 0.5$
	0.3	0.7	1

$P(A/B') = 0.4$

$\Rightarrow \dfrac{0.2}{x} = 0.4$

$\Rightarrow x = \dfrac{0.2}{0.4} = 0.5$

a) $P(A'/B') = \dfrac{0.3}{0.5} = \underline{0.6}$

Alternatively

$P(A/B') + P(A'/B') = 1$

$\Rightarrow P(A'/B') = 1 - P(A/B')$

$= 1 - 0.4$

$= 0.6$

b) ~~P(A/B)~~ $P(B/A) = \dfrac{0.1}{0.3} = \dfrac{1}{3} = \underline{0.\dot{3}}$

4.

a) From calc $\underline{r = -0.5089}$

b) Trace $\Rightarrow$ "less than 0.05" it looks as though Charles used $\underline{x = 0.05}$, as all other values given to 1 d.p.

c) Sensible for interpolation $\Rightarrow$ $0 \le x \le 19$

however $8.13 - 0.49x = 0$

$\Rightarrow \dfrac{8.13}{0.49} = x$

$\Rightarrow x = 16.6$ $\Rightarrow$ $\underline{0 \le x \le 16.6}$

as if $x > 16.6$ $y < 0$ which is not sensible

d) Relationship on graph $\underline{\text{looks non-linear}}$

1e) b negative as graph  not 

f) $y = ax^b \Rightarrow \log_{10} y = \log_{10} a + b \log_{10} x$

Both x and y data contain 0, can not find $\log_{10} 0$

5 a), b), c)  symmetrical data with mode in middle

6. Area under graph represents probabilities

Section 2

1 a) $H_0 : \rho = 0$

$H_1 : \rho > 0$

b)

	Size	r	%	Crit	Conclusion
i)	10	0.5	5	0.5494	Accept H_0
ii)	20	0.5	5	0.3783	Accept H_1
iii)	30	0.5	10	0.2407	Accept H_1
iv)	40	0.4	10	0.2070	Accept H_1

10

2.

	H_1	n	r	%	Crit region	Conclusion
a)	$\rho < 0$	n=20	-0.4	5	$r < -0.3783$	Accept H_1
b)	$\rho \neq 0$	n=30	0.5	2	$r > 0.4226$, $r < -0.4226$	Accept H_1
c)	$\rho \neq 0$	n=40	-0.3	1	$r > 0.4026$, $r < -0.4026$	Accept H_0

6

3. a) From calc $r = -0.3454$

b) i) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

n=10, 10% $\Rightarrow$ crit values = ± 0.5494

$\Rightarrow$ Accept H_0, Reject H_1

i.e Insufficient evidence that there is linear correlation between Daily Mean Pressure and Daily Mean Wind Speed

ii) $H_0 : \rho = 0$

$H_1 : \rho > 0$

$r < 0 \Rightarrow$ Accept H_0 i.e insufficient evidence that there is positive linear correlation between ...

iii) $H_0 : \rho = 0$

$H_1 : \rho < 0$

n=10. 10% $\Rightarrow$ crit value = -0.4428

$\Rightarrow$ Accept H_0, Reject H_1 i.e insufficient evidence that there in negative linear correlation between ... ...

c) knots

d) No, would get same value of r and same conclusions in any unit. Linear scaling does not affect r

14

4. $r = -0.4367$

5% $\rightarrow$ H_1 | H_0

a) $H_0 : \rho = 0$

$H_1 : \rho < 0$

n = ? 5% $\Rightarrow p = 0.05$

n=15 $\Rightarrow$ crit value = -0.4409

n=16 $\Rightarrow$ crit value = -0.4259

$\Rightarrow$ smallest n is 16

b) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

n=? 5% $\Rightarrow p = 0.025$

n = 20 $\Rightarrow$ crit values = ± 0.4438

n = 21 $\Rightarrow$ crit values = ± 0.4329

$\Rightarrow$ smallest n is 21

6

5. a) $r \approx 0.8$ ✓ strong positive linear correlation

b) $r \approx -0.8$ ✓ negative exponential linear correlation (r is misleading)

c) $r \approx -0.1$ ✓ 2 distinct sets of negative linear correlation (if treated as one set of data, hardly any correlation)

d) $r \approx 0.4$ ✓ 4 Very slight positive linear correlation (2 outlier values have big affect on r)

Total = $\boxed{40}$